10/539692

Rec'd PCT/PTO 16 JUN 2005

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

[Continued on next page]

(54) Title: ILLUMINATION INVARIANT FACE RECOGNITION
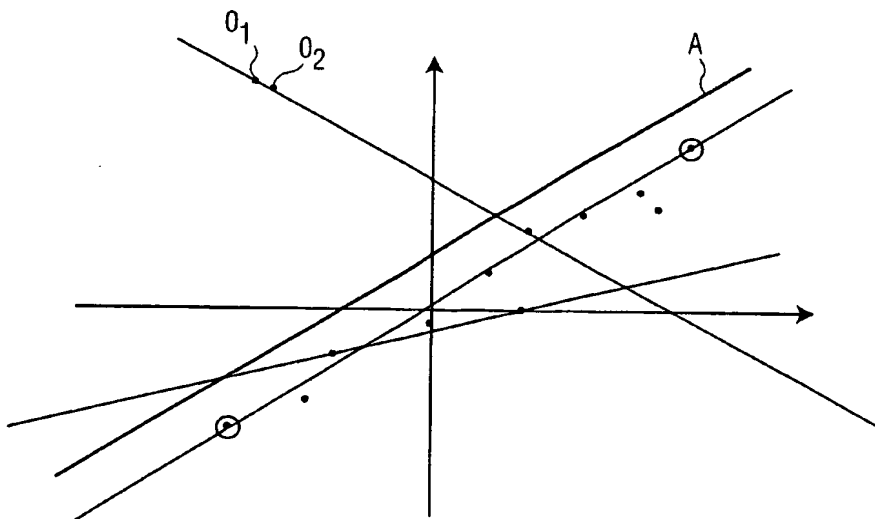
(57) Abstract: Random sample subsets are selected from an image. The RANSAC algorithm is used to perform robust matching between the samples of a captured image and like samples of stored images. By selecting enough samples the likelihood that all samples will be affected by lighting changes is less and therefore it is more likely that a match will be found even in the case of different lighting in the captured image than the stored images. Additionally a weighted sampling technique is used which selects samples based on likelihood of being affected by lighting changes.

WO 2004/057522 A1

WO 2004/057522 A1

ILLUMINATION INVARIANT FACE RECOGNITION

This invention relates in general to face recognition and in particular to light

invariant face recognition.

5   Face recognition systems are used for the identification and verification of

individuals for many different applications such as gaining entry to facilities, recognition

of people to personalize services such as in a home network environment, and locating

wanted individuals in public facilities. The ultimate goal in the design of any face

recognition system is to achieve the best possible classification (predictive) performance.

10   Depending on the use of the face recognition system it may be more or less important to

make sure that the comparison has a high degree of accuracy. In a high security

application such as identifying wanted individuals, it is very important that identification is

achieved regardless of minor changes in the captured image vs. the stored image.

The process of face recognition typically requires the capture of an image, or

15   multiple images, of a person, processing the image or images and then comparing the

image with stored images. If there is a positive match between the stored image and the

captured image the identity of the individual can either be found or verified. U.S. Patent

No.6,292,575 describes such a system and is hereby incorporated by reference.

The processing of the images includes normalization. Normalization insures that

20   the size of the face in the captured image is relatively the same size as in the stored images.

This normalization is an attempt at controlling differences that might occur in a captured

image which are not found in a stored image (and visa-versa). These differences cause

false negative identification results when the two faces are compared but when in actuality

the same person appears in both images.

Many systems also control the lighting of the captured image to insure that the

lighting will be similar to the lighting of the stored images. Once the individual is

positioned properly the camera takes a single or multiple pictures of the person and a

comparison is made to stored pictures. A problem with these systems is that for many

5   applications, such as for security applications, sometimes the lighting of the captured

image is different than the lighting used for the stored image. When the lighting is

different in the captured image than the lighting in the stored image there is a chance a

false negative identification result. Since it is not always possible to have the same

lighting, a system is needed which will take into account the different lighting.

10   Accordingly, it is an object of the invention to provide a system that performs face

identification/verification which is less susceptible to errors caused by light variation.

This object is achieved by taking random samples of pixels within the face and

performing comparisons with the same random samples in the stored images. The

comparison is performed using almost any form of face recognition. The RANSAC

15   algorithm is used to perform robust matching. The RANSAC algorithm helps eliminate

.the possibility of having shadowed regions of the face causing false negative results.

In one embodiment of the invention instead of a purely random sampling, a

weighted sampling is used which is weighted in the areas of the face which are least likely

to be affected by lighting. A determination is made as to which areas of the face are more

20   likely to be affected by lighting. The samplings are then weighted in the areas where there

is a lower probability of light changes. These weighted samples are then compared to the

like pixels in the stored images using face recognition to determine similarity.

Other objects and advantages will be obvious in light of the specification and

claims.

2

For a better understanding of the invention reference is made to the following

drawings:

Fig. 1a shows the selection of a line through a set of points in the presence of outliers.

Fig. 1b shows the selection of a line through a set of points in the presence of outliers using

5   the RANSAC algorithm.

Fig. 2 shows a flow chart of a preferred embodiment which uses the RANSAC algorithm

to perform robust face recognition.

Fig. 3a shows images of random pixel samples of size 1000, for an image of size 64x72.

Fig. 3b shows random pixel samples for real images of size 133x100, and the random

10   sample being 40% of the image.

Fig. 4 shows a flow chart of how weighted sampling is used in an exemplary embodiment

of the invention.

Fig. 5 shows a security system in accordance with a preferred embodiment of the

invention.

15   In typical face recognition programs pixels of a captured image are compared to

pixels of stored images using face recognition methods. If a certain percentage of pixels

match the same pixels in a stored image, the images are deemed to match and a verification

or identification has occurred. There are many known face recognition methods which

determine similarity between images. The term "match" as used herein can mean a

20   probability of a match, an exact match, or enough pixels match above a predetermined

threshold that identification is probable, or the similarity index found after performing face

recognition on these samples is above a certain threshold. The problem is that if a portion

of the face in the captured image has a shadow from a light source, and the image of this

person in the stored database is shadow-free, when the pixels of the two images are

3

compared it is likely that those pixels within the shadow will not match the corresponding

pixels in the stored image. This is because there will be a large percentage of non-

matching pixels due to uneven lighting rather than due to the faces in the images being

from different people. If a certain percentage of pixels must match to be deemed a proper

5   identification, the large percentage of non-matching pixels will cause a false negative

match.

It has been found that face recognition can be performed on portions of a face

rather than an entire face. The present invention chooses random samples of pixels within

the face but these random samples are chosen as part of an algorithm which ignores

10   "outliers", i.e. data points that lie outside of a majority of the data points. The RANSAC

algorithm is a robust estimation algorithm that avoids "outliers" from the chosen samples

from causing incorrect matching results and it is used in a preferred embodiment of the

instant invention. Other such algorithms can also be used such as the least median of

squares (LmedS). The RANSAC algorithm, as explained below in more detail, chooses

15   pixels randomly and detects "outliers" which in the present invention are the shadowed

regions.

In a preferred embodiment of the invention the captured image is compared to the

stored images using a known face recognition algorithm and if no match is found, then the

captured image is randomly sampled and the random sample of pixels T_j is compared to

20   the same random sample in the stored images using a face recognition algorithm. The

number of random samples selected is determined by the computing power and the degree

of accuracy that is needed for a particular application. In a preferred embodiment,

anywhere from 100 to 1000 or more sets of samples are chosen. Face recognition is

performed on all samples. The sample that provides the best result (i.e. the sample that

4

most closely matches the same sample in a stored image) is chosen and the result of the

sample comparison is compared to a threshold. If the sample of the captured image

matches the same sample in a stored image then the stored image is deemed to show the

same person as in the captured image. The stored image that provides this match has a

5   higher probability of being the same face as in the captured image since it is likely that one

of the samples will be relatively free of outliers. This results in fewer false negatives since

the random sample is less likely to include the portions of the face that are affected by

lighting.

Fig. 2 shows a flow chart of a preferred method of this invention. At 10 an image

10   is acquired. At 12, for I=1:X a random sample subset 20 is created. I is an increment and

X is determined based on computing power and degree of accuracy needed. The

RANSAC algorithm is used to select the subset and the number of subsets. RANSAC

stands for Random Sample Consensus and it will be explained herein in terms of line

estimation. It works similarly in other cases such as in images. Assume that a random

15   sample of points exists as shown in Fig. 1a, and a line A must be fitted through these

points. If there are outliers (there are two O_1 and O_2 in Fig. 1a, in the northwest quadrant)

the line A will be significantly shifted from the true position. Fig. 1b shows this same

random sample of points, but the line A is fitted using the RANSAC algorithm. In Fig. 1b

the line A is a better line estimation than Fig. 1b if we assume that O_1 and O_2 are erroneous

20   data points. RANSAC works the following way:

1.   Randomly pick a minimal number of points needed to estimate the line

(two).

2.   Fit the line through these points.

3. Count the number of points from the set that are close to the line (according to a given criteria, i.e. all the points that are closer than three pixels are OK –they support the model.) Denote the number N_j, which quantitatively represents the concentration of points near the line.

4. Repeat steps 1-3 enough times to ensure with a certain probability (usually 95-99%) that at least the two points that have been selected are a good representation of a concentration of the points.

5. Find the greatest of all N_j and the line that has produced this N_j. Using all N_j points that support this line, estimate the optimal line using ordinary least squares method.

To perform general random sampling of pixels in an image, we represent an image as a vector, i.e., each pixel is assigned a unique number as shown in Table 1.

64

1	2	3	...	63	64
65	66	...			128
...					
4481	4482	...			
4545	4546			4607	4608

72

Table 1. Pixel annotation for a 64×72 image.

For the example given in Table 1, to create a sample of 100 random pixels we create a sample of random numbers between 1 and 4616, and then just choose pixels corresponding to the 100 random numbers. Fig. 3a shows images of random samples of size 1000, for an image of size 64x72. Fig. 3b shows random pixel samples for real images of size 133x100, and the random sample being 40% of the image. There is a wide majority of algorithms to create random numbers available in any programming language (C/C++, Fortran) or programming package (MATLAB, Mathematica, etc).

In a preferred embodiment of the present invention, random sampling is used in the form of RANSAC which is applied to perform face recognition as follows:

1.   Randomly pick a certain number of pixels from a facial test image (T_j).

2. Compare T_j with the corresponding (location-wise) pixels of the face in the database and compute a similarity measure s_j. (Note that s_j corresponds to N_j in original RANSAC algorithm.)

3. Repeat steps 1-2 enough times X to ensure with a certain probability (usually 95-99%) that all of the sample points are not significantly affected by light changes. X can be estimated experimentally or computed as follows:

Assume that the facial image in question is formed of 1000 pixels and that 50 of them are affected by light changes. Further assume that only 100 facial pixels can be selected to perform face recognition. If all 100 pixels are from the 950 pixels that are not affected by light changes, then there should be good recognition results. The probability that none of the selected 100 pixels are affected by light changes is given by

$$p = \frac{\binom{950}{100}}{\binom{1000}{100}} = 0.0044758$$

If it is necessary to guarantee that at least one sample is outlier free with the probability of .99, then the number of samples X that needs to be chosen can be obtained from

$$(1-p)^X < .01 \Rightarrow X \geq \frac{\ln 0.01}{\ln(1-p)} = 1027$$

4. The measure of similarity is now obtained as $s = \max(s_j)$.

5. If s is above some given threshold, then we have a match, otherwise there is no match.

As explained above, the sets of regions are chosen randomly; for a large enough X it is likely that at least one set of regions will have consistent lighting that will match well with the corresponding set of regions in the database. If there is a high similarity, then

5   there is a match between the captured image and the stored image with high similarity.

Referring back to Fig. 2. after the image is acquired at 10, then for X = 1 to 1000 a random sample subsets (20) are chosen. For each sample, face recognition is performed on the sample by comparing using the RANSAC or other outlier detection algorithm, the sample of the captured image with the same random samples in the stored images at 30. If

10   there is a high similarity 32 between the two samples then assume a match 33 between faces. Store the subset with the highest number of matches and determine if the number is high enough to declare a match between the images. If not then there is no match 34. In a first embodiment the face recognition is performed until a match is found or until X is reached. In a second embodiment, all samples are compared and the sample with the

15   highest similarity is compared to a threshold to determine if there is a match.

WEIGHTED SAMPLING FOR LIGHT INVARIANT RECOGNITION

While in uniform sampling all the pixels have the same probability of being chosen, in weighted sampling, some pixels are more likely to be chosen than the others. Weighted

20   sampling has also been studied extensively in literature, and we will explain the difference between it and uniform sampling through the following example.

Let us assume that we need to choose one number between 1 and 5 (array X), but that sampling is not uniform but weighted, as shown in Table 2.

X	1	2	3	4	5
Weights	.1	.2	.3	.2	.2

25   Table 2

9

One way to perform weighted sampling using the weights set forth in Table 2 would be to form an auxiliary array Y ={1,2,2,3,3,3,4,4,5,5}

This array has 10 elements and, if uniform sampling is performed on the array Y it

5   is equivalent to a weighted sampling of X. In this example, 2 is two times as likely to occur as 1; and 3 is most likely to occur with the probability of .3.

Weighted sampling for light-invariant face recognition is simply an extension of the use of the RANSAC algorithm in face recognition. In the previous case, sampling is performed uniformly, i.e. each pixel of the set T_j is chosen with equal probability 1/(# of

10   pixels in the face). In weighted sampling some pixels are more likely to be chosen than the others based on the following criteria.

Assume there is a face that has been photographed for 100 different illuminations. Compare each pixel p_j in the original facial image (i.e. with uniform illumination) with the same pixel in the images obtained under different illuminations and count how many times

15   the difference between them is below some given threshold. Let us denote this number n_j, and let

$$N = \sum_j n_j$$

The pixels with higher n_j are more useful as they are less likely to be affected by illumination changes, so these pixels result in higher accuracy if they are part of the

20   comparisons. In other words, using these pixels one is more likely to find a comparison that is illumination invariant. Accordingly in the weighted-sampling embodiment of this invention, these "light invariant pixels" are chosen more frequently as part of T_j . The probability of a pixel being selected in the weighted sample is not uniform but is given by the formula

10

$$prob(j) = \frac{n_j}{N}$$

The algorithm now proceeds the same way as with the RANSAC illumination invariant

recognition described before.

Alternatively, to compute weights for sampling, computer graphics can be used. A

5   3D model of a head is created with light sources placed in different directions. Many

computer graphics packages can then turn the 3d image back into 2 dimensions with the

different light sources creating different effects on the 2 dimensional images. One can then

determine which areas of the face have a higher likelihood of changing with different light

sources. The portions of the face that have a low likelihood of changing can then receive a

10   higher weight or a higher probability of being chosen than the pixels with a high likelihood

of changing. Examples of programs that perform this type of ray tracing are POV-Ray, or

OpenGL.

Fig. 4 shows a flow chart of a method of this invention where a weighted

sampling is used. In this case an image is acquired at 10. The portions of the image most

15   affected by light are determined at 15. The pixels are then weighted at 35 in favor of the

light invariant pixels. At 12, for I = 1:X a weighted sample subset is chosen at 45. At 55

the weighted subset is compared using same form of face recognition with the same

weighted sample of the stored image. A probability of a match is then determined based

on the subset with the best result.

20   Fig. 5 shows a security application in accordance with the instant invention. An

image is captured by an imaging device 70 of a person 72 walking into a room with an

overhead light source 71. The image is compared to images stored in computer 73 in

accordance with the invention.

11

While there has been shown and described what is considered to be preferred embodiments of the invention, it will, of course, be understood that various modification and changes in form or detail could readily be made without departing from the spirit of the invention. It is therefore intended that the invention be not limited to the exact forms

5   described and illustrated, but should be constructed to cover all modifications that may fall within the scope of the appended claims.

10

15

CLAIMS:

1. A method of comparing a captured image to stored images to find a match, comprising:

   retrieving a captured image (10);

   setting default similarity index;

           for I=1 to X (12),

           take a random sample of pixels of the captured image (20);

           perform face recognition between the random sample of the pixels of the captured image to corresponding random samples of pixels of the stored images (30);

           calculate a new similarity index for the random sample of the captured image and replace the similarity index if new similarity index is higher (32.33,34);

           increment I; and

           determine if the similarity index is above a predetermined threshold (32,33,34) for a match.

2. The method as claimed in claim 1, wherein a robust algorithm {Fig. 1b(A)} is used which samples the image and performs face recognition by substantially removing outliers from having an impact on the comparison results.

3. The method as claimed in claim 2, wherein the algorithm is the RANSAC algorithm {Fig. 1b(A)}.

4. The method as claimed in claim 2, wherein the algorithm is the least medium of squares algorithm.

5   5. The method is accordance with claim 1, wherein a weighted sample is used which is weighted (35) with pixels which have a low likelihood of being affected by light.

6. The method in accordance with claim 1, wherein the step of determining occurs before the step of incrementing I, and if there is a match then additional random
10   samples do not need to be taken and compared.

7. The method in accordance with claim 1, wherein the value of X is based on desired accuracy.

15   8. The method in accordance with claim 1, wherein the value of X is based on computing requirements.

9. A device for comparing a captured image to stored images to find a match, comprising:

20   a device which receives a captured image (10);

        a processor (73) which performs the following function:

            create default similarity index;

            For I=1 to X (12),

            random sample the pixels of the captured image (20);

perform face recognition (30) between the random sample of pixels

of the captured image and the corresponding random sample of pixels of the stored

images;

calculate a new similarity index and replace similarity index with

5   new similarity index if it is higher; and

increment I;

determine if the similarity index is above a threshold for a match

(32,33,34).

10   10.   The device as claimed in claim 9, wherein the processor performs image

comparison using an algorithm that substantially lessens the possibility that

outliers will cause inaccurate results {Fig.1b (A)}.

11.   The device as claimed in claim 10, wherein the algorithm is the RANSAC

15   algorithm {Fig.1b(A)}.

12.   The device as claimed in claim 10, wherein the algorithm is the least medium of

squares algorithm.

20   13.   The device as claimed in claim 9, wherein a weighted (35) random sample is used

which is weighted with pixels which have a low likelihood of being affected by

light.

14. The device as claimed in claim 9, wherein the determining occurs before the step

of incrementing I, and if there is a match then additional random samples do not

need to be taken.

5   15. The device as claimed in claim 9, where the value of X is based on computing

power.

16. The device as claimed in claim 9, where the value of X is based on desired

accuracy.

10

17. A device for comparing images, comprising

a random sampler (10) for selecting a plurality of random samples of pixels

from a captured image;

a comparator (30) that performs face recognition on the random sample and

15   on a corresponding random sample in a stored image such that shadow regions of

the random samples are treated as outliers and do not substantially affect the

outcome of the face recognition;

a processor (32,33,34) that computes a similarity index and determines if

there is a match between the captured image and the stored image.

20

18. The device as claimed in claim 17, wherein a plurality of random samples (10)

are chosen and compared (30) and the random sample with the highest similarity

index (32,33,34) is used to determine if there is a match between the captured

image and the stored image.

16

19.   A method of comparing images, comprising:

        selecting a random sample (10) pixels from a captured image using a robust

sampling algorithm;

5   performing face recognition (30), using this algorithm which substantially

ignores shadowed regions of the image, on the random sample and the same

random sample in a stored image;

        computing a similarity index (32,33,34) using this algorithm and

determining if there is a match between the captured image and the stored image.

10

20.   The method as claimed in claim 19 wherein the step of selecting selects a

plurality of random samples (10) and the step of performing face recognition

compares (30) the plurality of random samples with the same random samples in

the stored images and computes similarity indexes and the highest similarity

15   index is used to determine if there is a match between the captured image and the
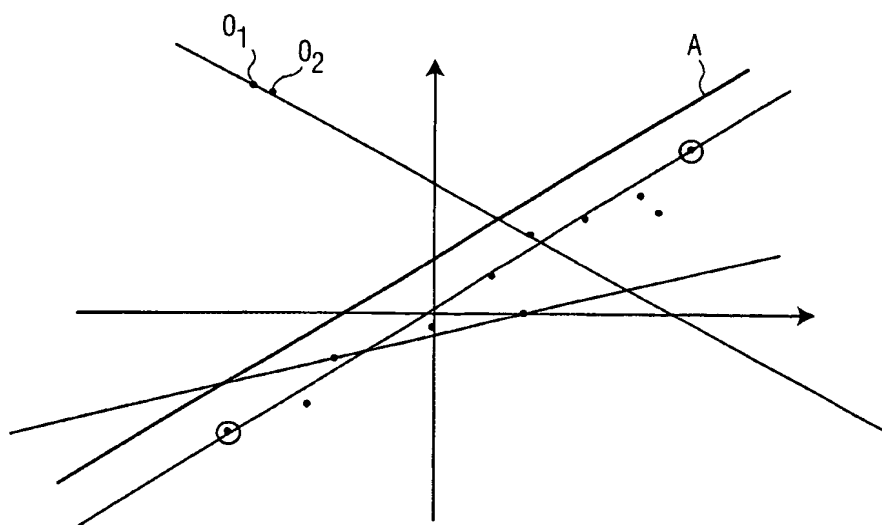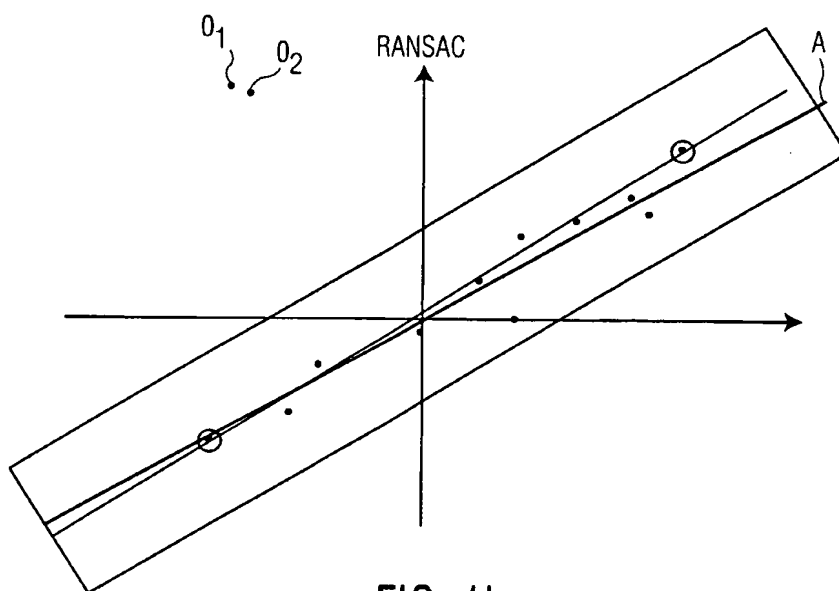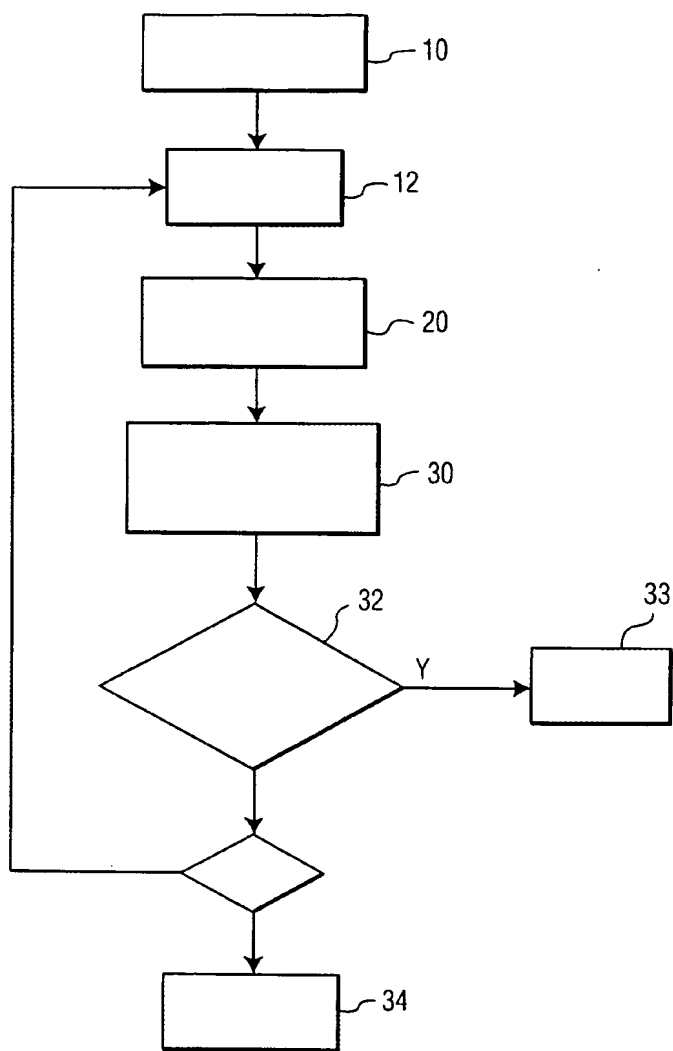
stored image.

20

FIG. 1a

FIG. 1b

FIG. 2

FIG. 3a

FIG. 3b
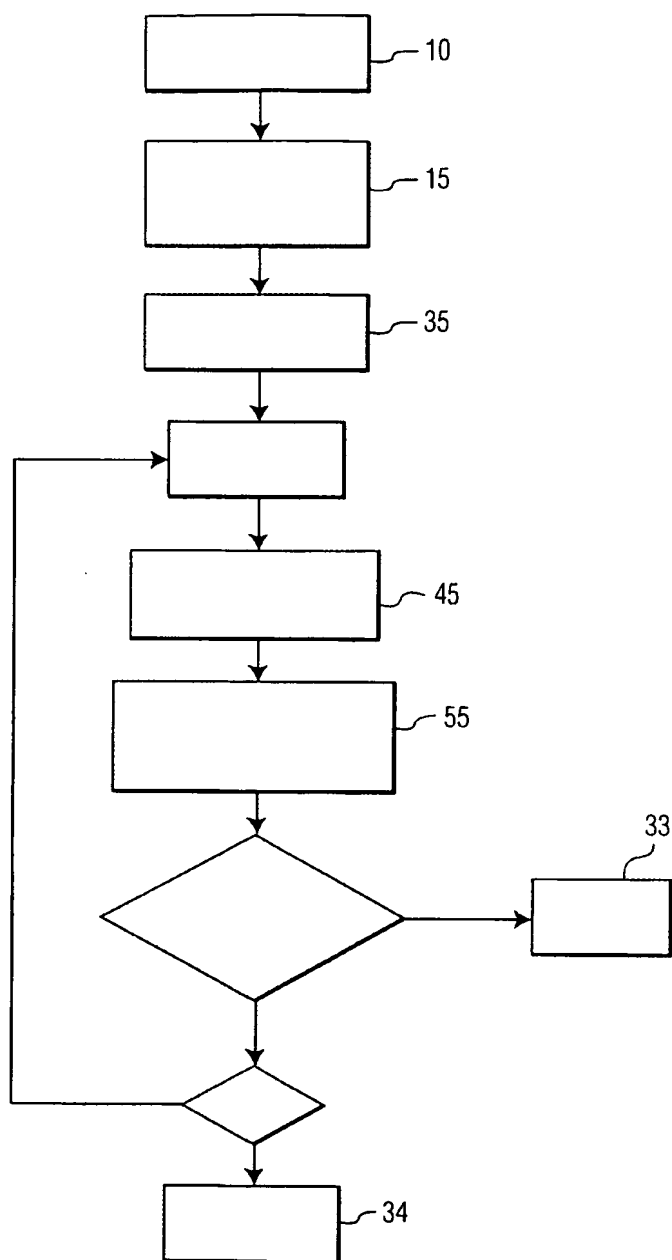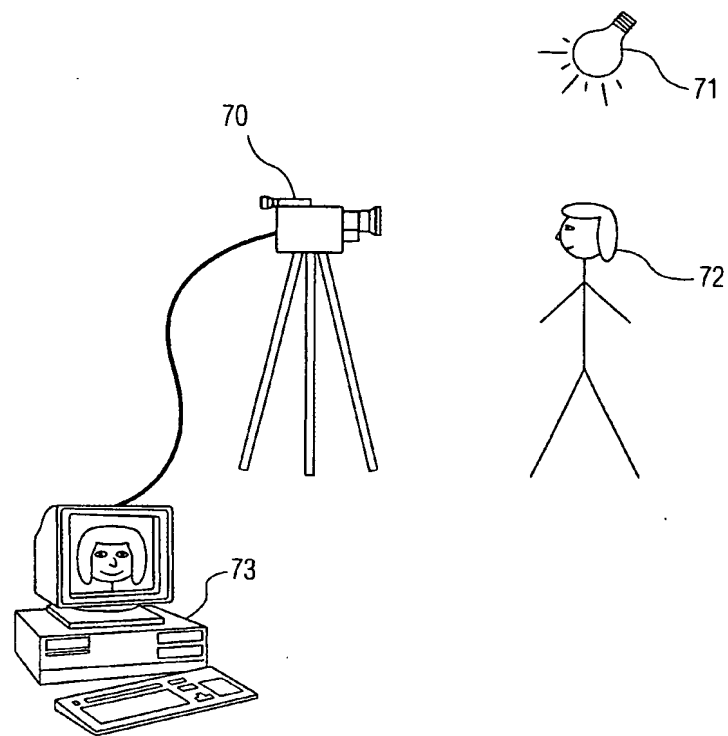
FIG. 4

FIG. 5

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 6072496	A	06-06-2000	AU	4550799 A	30-12-1999
			EP	1093616 A1	25-04-2001
			JP	2002517859 T	18-06-2002
			WO	9964961 A1	16-12-1999

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   G06K9/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7   G06K   G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 072 496 A (GRIMM CINDY MARIE ET AL) 6 June 2000 (2000-06-06) column 6, line 1 - line 67 column 9, line 52 - column 11, line 4 column 15, line 22 - column 17, line 63	1-20
A	DORNAIKA F ET AL: "Face model adaptation using robust matching and active appearance models" PROC. IEEE WORKSHOP ON APPLICATIONS OF COMPUTER VISION, 3 December 2002 (2002-12-03), pages 3-7, XP010628716 the whole document	1-20

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 May 2004	04/06/2004

| Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 | Authorized officer Grigorescu, C |

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	FISCHLER M A ET AL: "RANDOM SAMPLE CONSENSUS: A PARADIGM FOR MODEL FITTING WITH APPLICATIONS TO IMAGE ANALYSIS AND AUTOMATED CARTOGRAPHY" COMMUNICATIONS OF THE ASSOCIATION FOR COMPUTING MACHINERY, ASSOCIATION FOR COMPUTING MACHINERY. NEW YORK, US, vol. 24, no. 6, June 1981 (1981-06), pages 381-395, XP001149167 ISSN: 0001-0782 the whole document	1-20
A	ZHANG Z: "DETERMINING THE EPIPOLAR GEOMETRY AND ITS UNCERTAINTY: A REVIEW" INTERNATIONAL JOURNAL OF COMPUTER VISION, KLUWER ACADEMIC PUBLISHERS, NORWELL, US, vol. 27, no. 2, 1 March 1998 (1998-03-01), pages 161-195, XP000755415 ISSN: 0920-5691 the whole document	1-20